Combining Systems Thinking and Game Theory to Develop Alternative Strategies for Assessing University Bursary Physical Education in New Zealand Secondary Schools

Robert Y. Cavana
Senior Lecturer in Decision Sciences
Victoria Management School, Victoria University of Wellington,
PO Box 600, Wellington, New Zealand
Email: bob.cavana@vuw.ac.nz

Roger Moses
Principal, Wellington College,
PO Box 16-073, Wellington, New Zealand
Email: hmo@Wellington-College.school.nz

Abstract

This paper addresses the New Zealand Qualification Authority's (NZQA) current scaling practices for assessing University Bursary Physical Education (UBPE) in New Zealand secondary schools. The scaling procedures being used are based on the raw internally assessed marks provided by the secondary schools. Unfortunately, some schools appear to be providing 'inflated' raw marks for UBPE, thus engaging in 'game playing'. This has resulted in a great number of students and schools receiving final marks for UBPE at much lower levels than they deserve. This paper analyses the situation from Wellington College's perspective and suggests that 'systems thinking' and 'game theory' could be employed in 2003 by secondary schools to combat the ineffectiveness of NZQA in properly moderating UBPE assessments in the 'game playing' schools. However, while 'game theory' may provide a rational solution to the problem, it raises a number of ethical questions related to all stakeholders in the system.

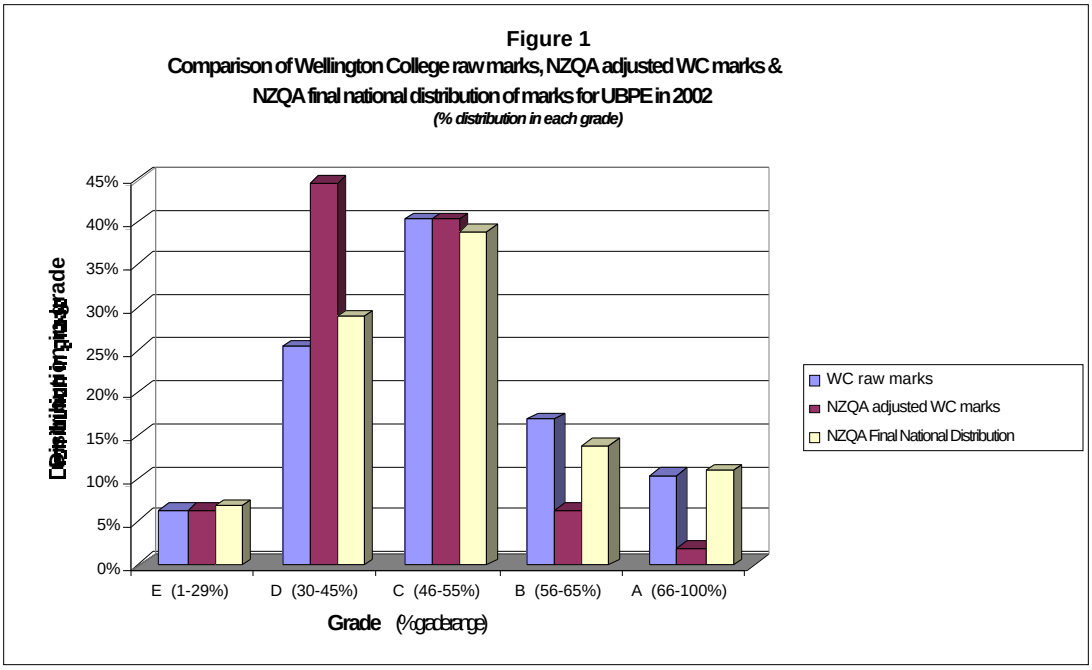
Keywords: NZQA, University Bursary Physical Education, Wellington College, assessment strategies, systems thinking, game theory

1. Introduction

In the 'University Entrance, Bursaries and Scholarships 2002 Candidate Information Sheet' supplied by the New Zealand Qualification Authority (NZQA), Forsyth (2003, p2) states that:

> *The internally assessed marks submitted by schools may be scaled. This is to ensure a fair and comparable standard between marks from different schools… The scaling of internally assessed marks prevents candidates being advantaged or disadvantaged by the particular school they attended.*
>
> *…Hence standards in different subjects have to be comparable. This prevents candidates being advantaged or disadvantaged by a particular choice of subjects.*

Unfortunately, the students studying University Bursary Physical Education (UBPE) at Wellington College (and similar schools) in 2002 were heavily 'disadvantaged', due to the NZQA's assessment procedures in this subject. This resulted in NZQA scaling down Wellington College's internally assessed raw marks for UBPE by an average of 11.7% (or 5.9 marks) from a mean of 50.4 (C, *a pass*) to 44.5 (D, *a fail*) out of 100. Also the students who received higher raw marks were scaled down more both in absolute and in relative terms.



Source: Wellington College and NZQA (2003b)

Figure 1 shows that based on the professional judgement of Wellington College PE teachers, 5 students (11%) should have received an A grade, and 8 students (17%) should have received a B grade. However, NZQA severely scaled these results down so that only 1 student (2%) at Wellington College received an A grade, and only a further 3 students (6%) received a B grade.

Nevertheless, the NZQA (2003b) final national distribution demonstrates that on average in a class of 47 students, 5 (or 11%) would receive an A grade and a further 7 (14%) would receive a B grade. In other words, 25% of the students in a typical 'average' New Zealand Physical Education class would receive an A or B grade. This compares with 28% as assessed by Wellington College UBPE teachers for their 2002 bursary class, and 8% as finally scaled down by NZQA for this class.

The implication of the NZQA scaling was that Wellington College either provided sub-standard education at bursary level Physical Education, or that the quality of Wellington College UBPE students was well below the national average (or a combination of both these). However, this is clearly not the case, since in recent years an NZQA moderator had been to Wellington College and confirmed that the standards were at the appropriate national level.

Although NZQA ranked Wellington College on the 25th percentile for UBPE in 2002 (222nd out of 295 schools (Colbert, 2003a)) the cohort of UBPE students at Wellington College in 2002 was about the same academically as the average UBPE class in New Zealand, hence the class should have been ranked on about the 50th percentile (or about 147th out of 295 schools). This suggests that about 25% of the secondary schools in New Zealand could have provided 'inflated' marks for UBPE in 2002 (we refer to this as 'game playing' in this paper).

There are a number of implications of the unfair scaling by NZQA of Wellington College's Physical Education bursary marks in 2002. These include:

- Two students missed out on receiving a B bursary and one student an A bursary as a result of this scaling, with 39% of the class receiving grades below what they deserved based on the professional judgement of Wellington College PE teachers.
- Instead of 32% of the class receiving a failed grade (D or E) for the course, NZQA scaling resulted in over half (51%) of the class failing UBPE.
- The downgrading from a C average grade to a D (*fail*) average grade for the UBPE class is also unfair to the PE teachers at Wellington College, as it undeservedly reflects badly on them.
- Students considering taking PE at Wellington College (and similar schools) in 2003 may have heard about the biased treatment by NZQA and may have been discouraged from taking PE. This will also reflect badly for New Zealand society, as the Government is currently trying to educate New Zealander's to lead healthier and better lives. This is in sharp contrast to the Hon Trevor Mallard, Minister of Education's comments in The Dominion Post ('Let's get physical Mallard tells kids', reported by Mulrooney, 2003):

  "The objective for me is to get them off their backsides and into activities that reduce the size of their backsides"

It was also particularly disturbing to see the trends in the difference between Wellington College's internally assessed UBPE mean mark and the lower final mean mark provided by NZQA (see Figures 2 and 3). This has increased from about 0.5 marks in 2000 to nearly 6 marks lower in 2002. If this trend continues, the gap could be even greater in 2003.

Figure 2
Wellington College UBPE Mean Marks: 2000 - 2002

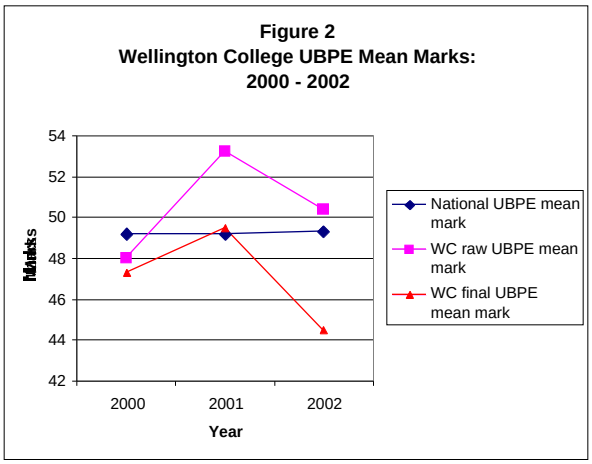
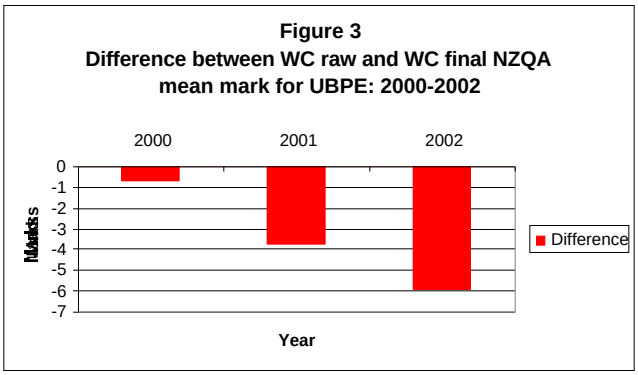


Figure 3
Difference between WC raw and WC final NZQA mean mark for UBPE: 2000-2002



Source: Wellington College & NZQA (2003b).

This paper summarises a recent report prepared by Cavana and Moses (2003) that was distributed by Wellington College to the Minister of Education, NZQA, and various other educational bodies in New Zealand. The next section provides a brief rationale for using systems thinking and game theory as analytical tools in this study. This will be followed by a summary of the actual procedures currently being used by NZQA for determining the final UBPE results. The next section briefly compares the performance of the Wellington College UBPE results with other Wellington College bursary results in 2002. Then the final marks for a selection of Auckland and Wellington secondary schools taking UBPE in 2002 will be compared. This analysis indicates that there appears to be some 'game playing' by some schools in New Zealand (ie schools providing 'inflated' marks for their internal assessment of UBPE above levels indicated by strictly adhering to NZQA guidelines). This leads us to the use of system thinking and game theory as a means of developing and evaluating alternative UBPE assessment strategies for Wellington College and similar schools. Finally some concluding comments and recommendations are provided.

2. Rationale for using Systems Thinking and Game Theory

From a systems thinking point of view, the trends outlined in Figures 2 and 3 represent a reference mode, which can be examined initially by developing causal loop diagrams. These can be used to help diagnose the problem, and to identify potential leverage points in the system for NZQA to intervene and correct the biases that have emerged. If NZQA fails to take appropriate corrective action in 2003, we suggest that Wellington College (and similar schools) then use game theory to help develop alternative assessment strategies for UBPE to combat the distortions caused by the 'game playing' secondary schools in New Zealand.

There do not appear to be many recent studies linking systems thinking / system dynamics with game theory. Rabbino (1998) has discussed applying Brandenburger and Nalebluff's (1996) principles of *Co-opetition* (cooperation and competition using game theory principles) with system dynamics tools; Kummerow & Quaddus (1998) & Kummerow (1999) have analysed office market cycles and policies with a system dynamics approach incorporating game theory; Gala et al. (1999) have undertaken a system dynamics study of the Spanish telecommunications market using differential game theory; and Warren (2002) has incorporated game theory into some of his strategy dynamic models.

However, most of these studies incorporate game theory into a system dynamics simulation model (eg see Forrester, 1961; Coyle, 1996; Sterman, 2000 or Maani & Cavana, 2000 for further details about system dynamics). This study is different in that it uses the problem structuring tools of system dynamics (ie causal loop diagrams) to help explain the reasons for the historical behaviour patterns for the UBPE marks, and then to show where the suitable intervention points are for NZQA to correct the situation. Since we are just concerned with the situation in 2003 at this stage, we do not develop a dynamic simulation model, but consider the choices for Wellington College (and similar schools) using a static game theory framework.

3. NZQA Scaling Procedures used by NZQA for PE Bursary Marks

The published information provided by NZQA does not fully explain how NZQA arrives at the final UBPE marks for each candidate and for each school. NZQA (2003a, p35) indicates that:

- "*UB Physical Education is fully internally assessed…*
- *Marks for Physical Education for a particular school are moderated by the final marks for the same group of candidates in all their other subjects. If the mean of the marks submitted by the school for Physical Education differs by more than 7 from the mean of the marks for the other subjects, the Physical Education marks are adjusted for that school…*"

Additional information was provided at a meeting at NZQA on 21 February 2003, where the actual scaling process for UBPE was outlined more fully. This included the following comments:

"*The UBPE final mark given to each student is based on their percentile ranking from the raw PE marks, applied to the list of 'other' UB subjects taken by the PE students. For example, a PE student who ranked on the 50th percentile on the raw PE scores will be assigned the mark on the 50th percentile of the other subjects taken by the PE students*."

The underlying problem is that NZQA does not adjust the raw UBPE marks submitted by each school to remove the inter-school biases (unless the mean of the raw marks submitted by the school for PE differs by more than plus or minus 7 marks (ie a range of 14 marks) from the mean of the marks for the other UB subjects taken by the PE students in that class).

If a school follows closely the assessment guidelines provided by NZQA, then their students will be ranked lower in the national averages for UBPE. Conversely, if a school is an 'easy' or 'generous' marker, then their students will be ranked much higher on the raw marks, and subsequently get higher final NZQA marks for UBPE overall.

NZQA claims that equity in UBPE marking is assured by their system of moderation, ie by going around the schools to ensure that UBPE teachers are marking at the appropriate levels. The NZQA personnel admitted to us that generally UBPE teachers around the country do NOT mark at the appropriate levels but mark higher. Hence it would seem that their system of moderation is not working.

4. Comparison with other Bursary Subjects at Wellington College

> *…Hence standards in different subjects have to be comparable. This prevents candidates being advantaged or disadvantaged by a particular choice of subjects.*
>
> (Forsyth, 2003, p2)

It appears that Wellington College PE students have been seriously disadvantaged by NZQA for choosing to study Physical Education at bursary level! Nearly every other bursary subject at Wellington College in 2002 received a average pass grade of A- B

or C, compared with NZQA's 'failed' average of D for Physical Education. Table 1 provides a comparison between the NZQA final scaled marks for PE and a range of other Bursary subjects at Wellington College. There is some similarity in the cohort of students studying PE and Geography, hence one would have expected that the mean mark for UBPE would have been about 50% or similar to the internally assessed mark, rather than the final NZQA mean mark of 44.5%.

Table 1
Comparison of NZQA Final Adjusted Bursary Marks at Wellington College in 2002

Bursary Subject	Mean of NZQA final adjusted marks		Number of WC students
	(%)	*Ave Grade*	*(No.)*
Physical Education	44.5	D	47
Geography	55.5	C	53
English	62.4	B	123
Maths with Statistics	54.1	C	136
History	57.3	B	48
Biology	61.3	B	46

Source: NZQA (2003b)

5. Comparison of Selected Auckland and Wellington Secondary School's PE Bursary Marks

The scaling of internally assessed marks prevents candidates being advantaged or disadvantaged by the particular school they attended.

(Forsyth, 2003, p2)

Although we are not suggesting which are the 'easy' marking schools, it is interesting to compare the final performance in UBPE in 2002 of a range of Auckland schools compared with a range of Wellington schools. Table 2 indicates that on average 10% of the students taking PE at the selected Auckland schools were awarded a scholarship compared with none at the Wellington schools. Overall nearly 42% of the students at these Auckland schools were awarded a scholarship, A or B grade in UBPE, compared with just 15% at the Wellington schools. Also the NZQA web site data indicates that the average final mark for UBPE in the Auckland schools was 56% compared with just 45.3% in the Wellington schools.

Table 2

University Entrance, Bursaries and Scholarships 2002 Physical Education Grade Distributions for Selected Auckland and Wellington Secondary Schools

(Percentage grade distribution)

	Total	Schol.	A	B	C	D	E
Auckland Grammar School	**72**	11.1	16.7	22.2	30.6	18.1	1.4
King's College	**29**	0	34.5	10.3	34.5	13.8	6.9
Rangitoto College	**56**	8.9	12.5	7.1	50	19.6	1.8
Takapuna Grammar School	**36**	16.7	13.9	16.7	22.2	19.4	11.1
Westlake Boys' High School	**37**	13.5	13.5	10.8	32.4	21.6	8.1
Ave Selected Auckland Schools	***46***	***10.0***	***18.2***	***13.4***	***33.9***	***18.5***	***5.9***
Hutt Valley High School	**25**	0	0	4	56	32	8
Rongotai College	**12**	0	0	0	66.7	33.3	0
St Patrick's College (Silverstream)	**35**	0	5.7	20	34.3	25.7	14.3
St Patrick's College (Wellington)	**17**	0	5.9	29.4	17.6	29.4	17.6
Wellington College	**47**	0	2.1	6.4	40.4	44.7	6.4
Ave Selected Wellington Schools	***27***	***0.0***	***2.7***	***12.0***	***43.0***	***33.0***	***9.3***
All New Zealand Schools	*3,962*	*2.5*	*8.5*	*14.2*	*39.3*	*28.5*	*6.9*

Source:New Zealand Qualifications Authority, web site:
http://www.nzqa.govt.nz/qualifications/ssq/statistics/school/, accessed 7 August 2003

The differences are further highlighted when a comparison of the percentile scores is done. For example the mark for the 75th percentile for the Auckland schools was 68%, which is higher than the mark for the 95th percentile for the Wellington schools! What this implies is that the top 25% of UBPE candidates in the Auckland schools have performed better than 95% of UBPE candidates from the Wellington schools. Although some difference does exist between the overall academic standards at the Auckland schools compared with the Wellington schools, this is most unlikely and suggests an element of 'game playing' by some Auckland schools, providing inflated raw marks for UBPE.

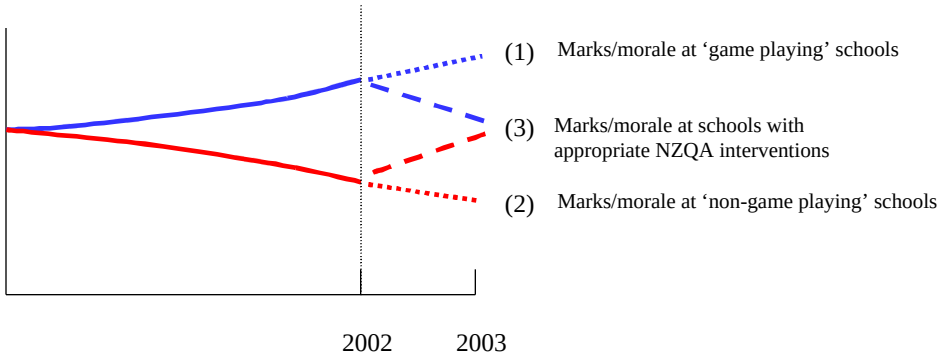
6. Systems Thinking Analysis of the PE Bursary Marks

The situation outlined in the previous sections indicates that some schools appear to be engaging in 'game playing' to inflate their internally assessed UBPE marks to higher levels than the levels indicated by the NZQA guidelines. Also many schools are still strictly adhering to the NZQA guidelines and are performing relatively worse in terms of their final NZQA marks for UBPE. As a consequence, the morale of the staff and students from these 'non-game playing' schools is also suffering and declining. This situation is reflected in Figure 4. Line (1) indicates the increasing UBPE marks and morale at the 'game playing' schools, whereas line (2) indicates the decreasing UBPE marks and morale at the 'non-game playing' schools. The unbroken lines represent the historical behaviour, whereas the small dotted lines represent the expected trend during 2003 if nothing different is done by NZQA (or the secondary schools collectively) to change this trend. The dashed lines (3) indicate a scenario whereby NZQA actively intervenes to moderate the school's internal assessments for UBPE much more carefully and consistently, or takes more appropriate corrective action to adjust inflated raw UBPE marks submitted by 'game playing' schools.

Figure 4

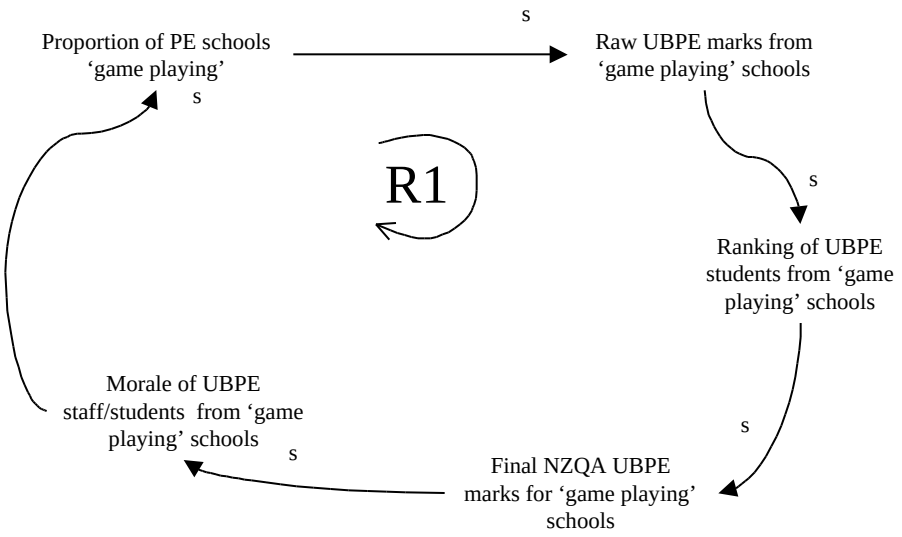
Behaviour over time chart comparing morale and UBPE marks

for 'game playing' and 'non-game playing' schools



We can now analyse the behaviour over time chart in Figure 4 above by utilising some of the systems thinking tools outlined in Maani and Cavana (2000). In particular, we will show how the causal loop diagrams in Figures 5 & 6 can help us analyse the situation.

Figure 5: Initial Reinforcing Feedback Loop (R1) for 'Game Playing' Schools



Key to Figs 5 & 6:

R1	Reinforcing loop – are positive feedback systems. They can represent growing or declining actions.
B1	Balancing loop – seeks stability or return to control, or aims for a specified target
o	Variable at the head of an arrow changes in the opposite direction to the variable at the tail.
s	Variable at the head of an arrow changes in the same direction as the variable at the tail.

We can analyse the variables and their relationships around the reinforcing feedback loop R1. We can start by observing that there appears to have been an increase in the number and proportion of secondary schools that have been involved in 'game playing' and submitting higher UBPE raw marks to NZQA. These higher raw marks result in higher rankings for the students of these 'game playing' schools and in higher final UPBE marks for these students, since the final NZQA rankings are based on the initial raw marks. This leads to an increase in the morale of teachers and students from these schools, which has a further reinforcing (increasing) effect on the proportion of schools involved in game playing, thus starting the cycle off again. We can now add the effects on the 'non-game playing' schools onto this diagram. This is shown in Figure 6 below.

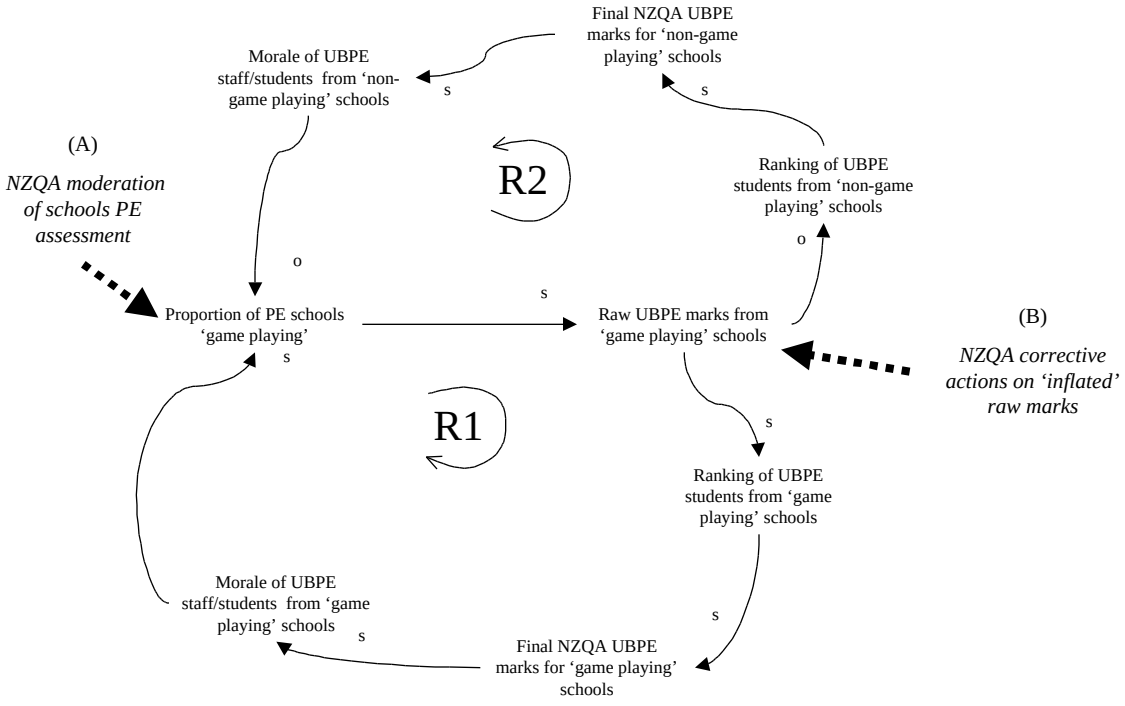
We have now added the extra variables and links that make up the second reinforcing feedback loop (R2) in this diagram. If we start our analysis again by considering that a higher proportion of 'game playing' schools will lead to higher raw UBPE marks submitted by these schools. However, because of the way NZQA ranks UBPE students on the basis on unadjusted raw UBPE marks, this will now result in lower rankings for UBPE students from 'non-game playing' schools, and hence lower final NZQA marks for these UBPE students and subsequently the morale of both the UBPE staff and students from these schools will suffer and decline. This will result in some schools changing their UBPE assessment strategies, and marking more generously the following year (eg in 2003), causing the first reinforcing loop (R1) to grow much more quickly. Hence this behaviour, based on the discussion of the reinforcing

Senge's 'Success to the Successful' systems archetype (Senge, 1990, pp385-386), whereby the 'game playing' schools are rewarded with higher final UBPE marks.

This behaviour, illustrated by lines (1) and (2) in Figure 4, can be averted by NZQA intervening at two leverage points in the system. The first point is indicated by point (A) in Figure 6, whereby NZQA can moderate more effectively the internal assessment of UBPE at all schools in New Zealand. The second leverage point is indicated by point (B). This indicates that if NZQA can not moderate the schools' internal assessment procedures appropriately, then it could intervene when it receives all the internally assessed raw UBPE marks by the schools, and make the appropriate corrections for bias then.

However, in recent years NZQA has consistently not intervened appropriately at the two leverage points identified in Figure 6. Hence we would recommend that 'non-game playing' schools might consider the implications of game theory discussed in the next section, and perhaps adjust their UBPE assessment strategies to overcome the bias caused by NZQA's failure to address the 'game playing' schools fairly.

Figure 6: Causal Loop Diagram for UBPE Assessments System



7. Assessing UBPE Performance using Game Theory

Unfortunately, with NZQA failing to adequately correct biases in the raw marks submitted by 'game playing' schools for UBPE, this now leaves 'non-game playing' schools in a position of having to 'guess' what the assessment strategies might be by the other schools. Fortunately a well developed body of literature called 'game theory' has been developed to help schools out in this situation.

The classic source for game theory is John von Neumann and Oskar Morgenstern's (1947) book, 'Theory of Games and Economic Behaviour'. Considerable literature is now available on this subject, including, for example Binmore (1990), Brandenburger and Nalebuff (1996), Romp (1997), Fisher & Waschnik (2002) and Watson (2002). However, we will use some of the ideas presented by Baumol (1965) here. For example, Baumol (p530) suggests that with game theory:

> *"It is possible to approach the analysis of competitive behaviour by a more deductive route. Instead of asking, inductively, what we can infer from the competitor's past behaviour, one seeks to determine a rival's most profitable counterstrategy to one's own "best" moves and to formulate the appropriate defensive measures."*

By competitor, we are now referring to the 'game playing' schools that offer UBPE, since it is by their 'assessment strategies' that Wellington College's relative ranking in UBPE has declined. This is evident from Figures 2 & 3, which show that the gap between Wellington College's internally assessed UBPE mean mark and the final mean mark provided by NZQA has increased significantly over the last few years. If this trend continues, the gap could be even greater in 2003, as more New Zealand schools adopt a more 'generous/inflated' marking strategy for UBPE.

The situation that Wellington College (and similar schools) face can be analysed as a 'two person, constant-sum' game (Baumol, 1965, p530). The final marks and percentile distribution for UBPE are based on the UBPE candidates performances in their 'other bursary' subjects, but the actual percentile for each UBPE candidate (and mean for the school) depends on their ranking based on the raw UBPE marks submitted by each school. Hence each school can vary their strategy on how they mark. We will define three illustrative strategies how each school might internally assess their UBPE marks:

Strategy 1	*Assess strictly according to NZQA guidelines*
Strategy 2	*Assess 'more generously' than NZQA guidelines (ie provide an average of about 5 more marks for the UBPE class than Strategy 1)*
Strategy 3	*Assess 'very liberally' compared to NZQA guidelines (ie provide an average of about 10 more marks for the UBPE class than Strategy 1)*

Now to help us with this analysis, we can regard the 'game' as being played by two players: Wellington College and all the other 'game playing' schools combined taking UBPE in 2003 (called 'Other Schools' here). For each strategy taken by Wellington College, a 'payoff' can be calculated depending on the strategy selected by the Other Schools. The payoff is measured by an indicative mean final mark for UBPE for Wellington College. The strategies and payoffs are summarised in the Payoff Matrix

in Table 3 below. About 25% of secondary schools in New Zealand offering UBPE are estimated to be engaged in 'game playing' behaviour.

Table 3
Payoff Matrix for Evaluating Wellington College's UBPE Assessment Strategies
(mean percentage mark for WC UBPE)

		Other Schools' Strategy		
		1	2	3
Wellington College Strategy	1	48%	44%	41%*
	2	52%	48%	44%*
	3	55%_	52%_	48%*_

In practice any number of strategies are possible, but we have simplified them to 3 strategies each in Table 3 to illustrate the analysis. For example in the year 2000, Wellington College employed strategy 1 in strictly assessing UBPE according to NZQA guidelines. By and large most of the Other Schools also appeared to be following Strategy 1 and Wellington College ended up with a final mean UBPE mark of nearly 48%, similar to its internally assessed raw mean mark (see Figure 2). However, by 2002 Wellington College was still assessing strictly to NZQA guidelines (ie strategy 1) but meanwhile the Other Schools appeared to be on average marking 'more generously' (ie strategy 2) and Wellington College received a final mean UBPE mark of about 44%.

As can be seen in Table 3, if both Wellington College and the Other Schools adopt the same strategies, then their payoffs will be the same, ie the average mean UBPE final mark for Wellington College will be 48%. For example, if both Wellington College and the Other Schools adopt 'very liberal' assessment strategies (strategy 3), then their students' relative rankings on their raw UBPE marks will be the same as if they both adopted very strict (strategy 1) or 'more generous' assessment strategies (strategy 2). In some respect this would be the 'fairest' outcome if NZQA does not take corrective action for 'game playing' schools in 2003. That is, if all NZ secondary schools offering UBPE in 2003 agree to follow one of the strategies outlined here!

The other payoff values shown in Table 3 have been estimated based on the strategies employed by each 'player'. These should be regarded as only indicative mean final UBPE marks for Wellington College.

The lowest payoff is estimated to be 41%, 7 marks below the 48% (the average of the 'other bursary' subjects taken by Wellington College PE students in 2002 who took 3 or more bursary subjects), and the highest payoff is given by a mean UBPE final mark of 55%, or 7 marks above the average mark for the 'other UB subjects'. In other words, there is a range of 14 marks that Wellington College could receive final UBPE mean marks from NZQA, depending on which strategies Wellington College and the Other Schools take! This is an immense variation currently outside the 'control' of NZQA.

7.1 Maximin and Minimax Strategies

Following Baumol (1965, pp530-534), we can now evaluate this payoff matrix to determine if there are optimal strategies, which both Wellington College and the Other Schools can take. We will also make the assumption that both players have full information about each other's strategies and can play any of the strategies in total. We will now assess the *maximin* and *minimax* strategies, since "the cautious approach to this problem is to assume the worst and act accordingly" (Baumol, 1965, p531).

Firstly, we will examine Wellington College's *maximin* strategy. For each strategy taken by Wellington College we will assume that the Other Schools adopt the strategy that will maximise their mean final UBPE mark. For example, if Wellington College adopts strategy 1, then we assume that the Other Schools will adopt strategy 3, since Wellington College will end up with a mean PE mark (41%) below the average and the Other Schools will end up with a mean UBPE mark above the average of 48% (the actual mark does not matter here and it will depend on the number of UBPE candidates in the Other Schools relative to the number of candidates in Wellington College.). We note this mark with an asterisk (*) to indicate the optimal strategy for the Other Schools. Similarly when Wellington College employs strategy 2, the Other Schools will also select their strategy 3, since it will minimise Wellington College's mean PE mark, and thereby maximise their own students' marks. This value (44%) is noted with an asterisk. So is the value of 48% when Wellington College selects strategy 3, the Other Schools will select their strategy 3 also. Now the optimal *maximin* strategy for Wellington College is to select the strategy that will maximise its average UBPE mark, ie "choose that one among its strategies for which the starred figure is highest…the maximum among these minimal payoffs" (Baumol, 1965, p532). That will occur when Wellington College employs strategy 3, ending up with a mean final UBPE average mark of about 48%. This is called their *maximin* strategy (ie the strategy that maximises the minimum payoffs for each of Wellington College's strategies).

The Other Schools can employ a similar approach (evaluating their *minimax* strategy). The worst for them is when Wellington College receives a high mean mark for UBPE, since that means they will receive a relatively lower mark. Now if the Other Schools adopt strategy 1, then the worst 'payoff' to them is if Wellington College adopts strategy 3 and receives a mean mark of 55%. This figure is marked with a prime. Similarly if the Other Schools play strategies 2 and 3 then the worst payoffs in each column (or highest mean marks for Wellington College) are marked with a prime also. The best of these pessimistic payoffs for the Other Schools is the lowest of these average Wellington College mean UBPE marks, which is 48%. Hence the best choice

for the Other Schools is strategy 3, which is called their *minimax* strategy (ie the strategy that minimises the maximum payoffs for Wellington College under each of the Other Schools strategies).

Now we can compare the *maximin* strategy for Wellington College, with the *minimax* strategy for the Other Schools to see if a 'saddle point' or 'equilibrium point' exists. In this case Wellington College's average mean UBPE mark of 48% from its *maximin* strategy 3 is exactly the same as the mean mark that the Other Schools expect Wellington College to receive when the Other Schools employ strategy 3 (ie the starred mark in row 3 coincides with the primed mark in column 3). Hence an equilibrium point does exist, and both Wellington College and the Other Schools would be best off employing strategy 3 (ie a 'very liberal' assessment of UBPE raw marks compared with NZQA guidelines).

7.2 Decision Making Under Uncertainty

"Decision making under uncertainty requires that the decision maker use his judgement and experience about future events" (Bierman et al. 1969, p9). We can now extend the analysis described above to incorporate some of the ideas from decision theory (eg see Baumol, 1965; and Bierman et al. 1969). In this situation the strategies employed by the Other Schools are considered 'states of nature' or 'environmental states' outside of the influence of Wellington College, and also independent of the actions of Wellington College. In some respects this is more realistic than the game considered above whereby Wellington College employs its *minimax* strategy 3, and the Other Schools respond by selecting their *maximin* strategy 3, thereby reaching an equilibrium point. In this analysis we assume that the Other Schools will employ their strategies irrespective of the strategies that Wellington College employs. However, we can attach probabilities to the likelihood of the Other Schools employing a specific strategy. In this case we assume that there is only a 5% chance that the Other Schools will return to strategy 1 (strictly following NZQA guidelines); an 80% chance that they will follow strategy 2 (more generous assessments) and a 15% chance that they will follow strategy 3 (very liberal assessment procedures). Sensitivity analysis can be undertaken to vary these probabilities.

This additional information is incorporated into the extended payoff matrix in Table 4. The expected value (or weighted average) for each of Wellington College's strategies can be calculated as follows:

Expected value of WC strategy 1 = 0.05 * 48% + 0.80 * 44% + 0.15 * 41%
= 43.8%

Similarly, the expected values for Wellington College's strategy 2 and 3 are 47.6% and 51.6% respectively. Now Wellington College can select the strategy that provides the highest expected mean mark for UBPE. This occurs for strategy 3.

Hence using this 'expected value' decision criterion, Wellington College would be encouraged to employ 'very liberal' assessment procedures compared to NZQA

guidelines (ie strategy 3). This analysis supports the decision reached above under the analysis of *maximin* and *minimax* strategies.

Table 4
Payoff Matrix for Evaluating Wellington College's UBPE Assessment Strategies under Uncertainty
(mean percentage mark for WC UBPE)

		Other Schools' Strategy *(environmental states)*			Expected WC mean UBPE mark
		1	2	3	
Wellington College Strategy	1	48%	44%	41%	43.8%
	2	52%	48%	44%	47.6%
	3	55%	52%	48%	51.6%*
Probability		*0.05*	*0.80*	*0.15*	

However, while this 'rational analysis' indicates that Wellington College should pursue its interests by following strategy 3, this raises some very important ethical issues:

(1) Should schools be involved in 'game playing' to ensure that their students and staff get fair assessments under the present NZQA UBPE assessment procedures?
(2) Is this 'game playing' ethical?
(3) What happens to the students that fail Bursary or get lower marks than they deserve in UBPE, because NZQA has failed to moderate the schools internal assessment properly, or has failed to take the appropriate corrective action when they have learned that some schools are not following NZQA guidelines as strictly as they should?
(4) Is NZQA's current behaviour ethical?
(5) What about the short and long term effects on the disadvantaged students?
(6) What happens to the schools that don't understand a 'game' is being played with the internal assessment of UBPE by an increasing number of schools? Will they be increasingly more disadvantaged in the 2003 UBPE final assessments by NZQA?

8. Conclusions & Recommendations

University Bursary Physical Education at secondary schools in New Zealand is a demanding and challenging subject, which requires a high level of physical and intellectual skills and ability to be able to perform well in it. It is also regarded as a significant university subject, with Physical Education graduates from New Zealand universities performing many valuable roles in the wider New Zealand society and in the international community.

There appears to be no reasonable justification for NZQA's excessive scaling down of Wellington College's University Bursary Physical Education marks in 2002, from an average of 50.4 (C) to a *'failed'* average of 44.5 (D). A case was made to NZQA (Cavana, 2003a; Moses, 2003a, b & c) and the Minister of Education (Cavana, 2003a, b & c) to have these marks scaled back up to the average of the marks submitted by Wellington College's Physical Education teachers, which were at similar levels to the final NZQA national distribution of bursary marks. However, this case was rejected by NZQA (Colbert, 2003a & b) and by Hon Trevor Mallard (2003a & b), the Minister of Education.

This paper has further investigated the situation and has applied systems thinking and game theory to demonstrate how Wellington College (or similar schools) can employ different assessment strategies to counterbalance NZQA's failure to correct for the higher marks generated by the 'game playing' schools in New Zealand. Nevertheless, while Wellington College and similar schools can also adopt 'game playing' to improve their marks and morale in UBPE in 2003, this raises a number of ethical issues related to all stakeholders in the system (ie the schools, UBPE candidates and staff, and NZQA).

Based on this analysis, we recommend that NZQA review their 2003 procedures for moderating the internal school assessment of UBPE and the ranking of unadjusted raw UBPE marks as the basis for the assessment of the final UBPE marks. However, if NZQA will not take the appropriate corrective action in 2003 against the 'game playing' schools, then the fairest outcome would occur if all schools that offer UBPE in New Zealand in 2003 adopt the same assessment strategy (ie assess UBPE compared with NZQA guidelines either 'strictly'; 'more generously'; or 'very liberally'). This will ensure that the outcomes in 2003 for UBPE for students and staff are 'relatively fair' and the distortions and biases from 'game playing' schools will be removed as much as possible.

The Minister of Education has advised us that: "From 2004 the University Bursaries award will be replaced with NCEA. Under this award all school assessment decisions will stand, avoiding the need for any statistical moderation" (Mallard, 2003b, p2). However, this would appear to actually increase the value of this research, because from 2004 onwards there will be a larger internally assessed component in all final year subjects at secondary schools in New Zealand.

Finally we would like to suggest that this paper contributes to the systems thinking literature, by demonstrating how systems thinking and causal loop diagrams can be used in conjunction with game theory to resolve important issues of concern to managers and other stakeholders in situations involving risk and uncertainty.

Acknowledgements

We would like to acknowledge the assistance of all the people who have helped in the preparation of this paper. In particular we would like to thank Messrs Robert Anderson, Chris Pollock, Chris Wells and Mrs Heather Oldershaw at Wellington College; Professors Cedric Hall and John Prebble at Victoria University of Wellington; various staff members at the New Zealand Qualifications Authority for all their assistance; and the anonymous ANZSYS referees for their comments. However, the opinions and interpretations in this paper remain the authors' own.

References

Baumol, W.J., 1965, *Economic Theory and Operations Analysis*, 2nd ed. Prentice Hall, Englewood Cliffs, NJ.

Bierman, H. Jr, Bonini, C.P, and Hausman, W.H., 1969, *Quantitative Analysis for Business Decisions*, 3rd ed. Richard D. Irwin, Inc., Homewood, Illinois.

Binmore, K.G., 1990, *Essays on the Foundations of Game Theory*. Blackwell, Cambridge, Ma.

Brandenburger, A. and Nalebuff, B.J., 1996, *Co-opetition*. Doubleday, New York.

Cavana, R.Y., 2003a, *Analysis of NZQA's Scaled Bursary Marks for Physical Education at Wellington College*. Victoria University of Wellington, Wellington. (Unpublished Report dated 3 February 2003, prepared for Mr Roger Moses, Principal, Wellington College).

Cavana, R.Y., 2003b, Letter to Hon Trevor Mallard dated 11 February 2003, re University Bursary Marks for Physical Education at Wellington College.

Cavana, R.Y., 2003c, Letter to Hon Trevor Mallard dated 25 February 2003, re University Bursary Marks for Physical Education at Wellington College.

Cavana, R.Y. and Moses, R., 2003, *Alternative Strategies for Assessing University Bursary Physical Education in New Zealand Secondary Schools*. Unpublished report, distributed by Wellington College, Wellington.

Cavana, R.Y., Delahaye, B.L., and Sekaran, U., 2001, *Applied Business Research: Qualitative and Quantitative Methods*. Wiley, Brisbane.

Colbert, K., 2003a, Letter to Roger Moses, Principal Wellington College dated 10 March 2003. Group Manager, Secondary Education Group, New Zealand Qualifications Authority, Wellington.

Colbert, K., 2003b, Letter to Roger Moses, Principal Wellington College dated 15 May 2003. Group Manager, NCEA Implementation and Secondary Education, New Zealand Qualifications Authority, Wellington.

Coyle RG., 1996, *System Dynamics Modelling: A Practical Approach*. Chapman & Hall, London.

Fisher, T.C.G. and Waschik, R.G., 2002, *Managerial Economics: a Game Theoretic Approach*. Routledge, London.

Forrester, J.W., 1961, *Industrial Dynamics*. MIT Press, Cambridge, Mass.

Forsyth, L., 2003, *University Entrance, Bursaries and Scholarsips 2002: Candidate Information Sheet*. New Zealand Qualification Authority, Wellington, January 2003.

Gala, M., Vicente, S.M., Gordillo, F., Alcala, I. and Aracel, J., 1999, Competition in the Spanish Telecommunications Market: a Co-evolutionary, Differential Game

Approach. *Proceedings of the 17th International Conference of the System Dynamics Society*, Wellington, July 20-23.

Kummero, M. and Quaddus, M., 1998, Office Market Cycles: a System Dynamics Approach to Improve Allocative Efficiency. *Proceedings of the 16th International Conference of the System Dynamics Society*, Quebec, July 20-23.

Kummero, M., 1999, Designing and Testing Policies to Prevent Office Building Oversupply Cycles. *Proceedings of the 17th International Conference of the System Dynamics Society*, Wellington, July 20-23.

Maani, K.E. and Cavana, R.Y., 2000, *Systems Thinking and Modelling: Understanding Change and Complexity*. Prentice Hall, Auckland.

Mallard, Hon Trevor, 2003a, Letter to R.Y.Cavana dated 18 March 2003. Minister of Education, Parliament Buildings, Wellington.

Mallard, Hon Trevor, 2003b, Letter to R.Y.Cavana dated 26 March 2003. Minister of Education, Parliament Buildings, Wellington.

Moses, R., 2003a, Letter to K. Colbert, New Zealand Qualifications Authority, Wellington dated 25 February 2003. Principal, Wellington College, Wellington.

Moses, R., 2003b, Letter to K. Colbert, New Zealand Qualifications Authority, Wellington dated10 March 2003. Principal, Wellington College, Wellington.

Moses, R., 2003c, Letter to K. Colbert, New Zealand Qualifications Authority, Wellington dated 13 May 2003. Principal, Wellington College, Wellington.

Mulrooney, P., 2003, Let's get physical Mallard tells kids. *The Dominion Post*, Wellington. Thursday, February 6, 2003, pA3.

NZQA 2003a, *2003 REFERENCE GUIDE: Sixth Form Certificate (transitional) (SFC(t)), Higher School Certificate (HSC), University Entrance, University Entrance, Bursaries and Scholarships*. Secondary Education Group, New Zealand Qualifications Authority, P O Box 160, Wellington.

NZQA, 2003b, University *Entrance, Bursaries and Scholarship Statistics 2002*. Percentage grade distribution by subject & Final score percentile statistics. New Zealand Qualifications Authority, Wellington. Web site accessed 7 August 2003: http://www.nzqa.govt.nz/qualifications/ssq/school/

Rabbino, H., 1998, Applying the Principles of Co-opetition with System Dynamics Tools. *Proceedings of the 16th International Conference of the System Dynamics Society*, Quebec, July 20-23.

Romp, G., 1997, *Game Theory: Introduction and Applications*. Oxford University Press, Oxford.

Senge, P.M., 1990, *The Fifth Discipline: The Art and Practice of the Leaning Organization*. Currency Doubleday, New York (First Currency Paperback Edition, October 1994).

Sterman, J.D. 2000, *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Irwin McGraw-Hill, Boston.

von Neumann, J. and Morgenstern, O., 1947, *Theory of Games and Economic Behaviour* 2nd ed. Princeton University Press, Princeton, N.J.

Warren, K., 2002, *Competitive Strategy Dynamics*. Wiley, Chichester.

Watson, J. 2002, *Strategy: an Introduction to Game Theory*. W.W. Norton, New York.